

Commentary

Robert J. Tetlow

Robert Arnold (2009) clearly and completely lays out the approach used by the Congressional Budget Office (CBO) for measuring potential output and discusses the limitations therein.

In this commentary, I revisit arguments made by the authors and discussants of a paper on this same subject at a 1978 Carnegie-Rochester conference to show how little the CBO methodology differs from methods used 30 years ago and conjecture on why this is so. From there I speculate on why current methods have been impervious to the critiques from 30 years ago and econometric developments in the years thereafter.

The measurement of potential output clearly matters, and matters even more in real time, at least for some decisionmakers. The *growth rate* of potential pins down the tax base for fiscal authorities and lawmakers; it provides a baseline for GDP growth for economic forecasters; and it helps establish a benchmark for policymakers and financial market participants to interpret the real-time data.¹ The *level* of potential defines the point to which the economy is expected to gravitate over the medium term and so is important for monetary authorities, forecasters, and anyone who needs to interpret business cycles. I review why and for whom it matters and critique the methods used by the CBO. The CBO methodology is not unique to that institution; rather, it is my impression that

a number of other, large macroeconomic forecast teams around the world use broadly similar tools. To the extent this is true, this critique is germane to a broader set of model builders and users.

After I provide some background, my comments get more specific. I argue that issues of econometric identification limit the confidence with which we can approach the CBO estimates; I argue against the widespread use of deterministic time trends, particularly in the real-time context; and I question the uncritical application of Okun's law.

WHITHER POTENTIAL?

Who needs potential output measures and for what reason? One way of illustrating this question from the perspective of a policymaker is to refer to a simple forecast-based Taylor rule, like the one shown below:

$$(1) \quad R_t = \underbrace{rr^*}_{\partial rr^*/\partial \Delta y^* \geq 1} + \phi_\pi \cdot \underbrace{E_t \pi_{t+1}}_{\partial E\pi/\partial y^* < 0} + \phi_y \cdot \underbrace{(y_t - y_t^*)}_{\partial(y - y^*)/\partial y^* < 0} + u_t,$$

¹ An example of the latter is the recent decline in labor force participation. Until a few years ago, sustainable employment growth from the establishment survey was estimated at around 120,000 and levels below that would have been interpreted as foreshadowing possible easing in monetary policy and increases in bond prices. The work of Aaronson et al. (2006) showed that sustainable additions to employment are probably much lower now than before.

Robert J. Tetlow is a senior economist in the Division of Research and Statistics at the Federal Reserve Board. The original discussion slides from the conference are available at the online version of this *Review* article. These slides—but not this text—use Federal Reserve Board/U.S. model vintages and associated databases and *Greenbook* forecast and historical databases (the latter of which under Federal Reserve Board rules are permissible for use only for forecasts currently dated before December 2002). The author thanks Peter Tulip, Dave Reifscheider, and Joyce Zickler for useful comments and Trevor Davis for help with the presentation slides.

Federal Reserve Bank of St. Louis *Review*, July/August 2009, 91(4), pp. 291-96.

© 2009, The Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

where R is the nominal federal funds rate, rr is the real funds rate, π is inflation, y is (the natural logarithm of) real gross domestic product (GDP), and u is a stochastic term. The asterisks on the real rate and on output represent “potential” (or “natural”) levels; these natural levels are not observable. The coefficients, ϕ_j , $j = y$, and π , would normally be expected to be positive. The partial derivatives above the equation itself show how changes in potential output affect the rule and hence decisionmaking. Starting with the term farthest to the right, an increase in the level of potential—that is, $\partial y^* > 0$ —decreases estimates of the output gap, $y - y^*$, all else equal. Higher potential would also reduce expected future inflation— $E_t \pi_{t+1}$ —because smaller gaps usually mean less inflation and both of these would be expected to lead to a lower federal funds rate. An increase in the growth rate of potential— $\partial(\Delta y^*) > 0$ —raises the equilibrium real interest rate, which would call for an increase in the funds rate, all else equal, but it would also have complex, model-dependent effects on current and future output gaps and inflation.²

What complicates this is that the only observables in the equation are current output, which is subject to revision, and the federal funds rate itself. A policymaker—in this instance, the Fed—is obliged to add structure to this underidentified equation through the use of a macroeconomic model of some sort. For their part, interpreters of the data—Fed watchers, among others—are obliged to “invert” the (perceived) policy rule and infer what the Fed’s estimates of rr^* , Δy^* , y^* , and $E_t \pi_{t+1}$ might be.³ The only inevitability is that *all* parties will get it wrong; the question is in what way and how critically.⁴

² I am thinking of a closed economy here, or at least one that, if open, is not “small.”

³ Of course, what Fed watchers might also want to infer from policy decisions given a policy rule is an estimate of the target rate of inflation. The target rate has been normalized out of our policy rule, for simplicity.

⁴ It makes a difference whether it is the Fed that is “getting it wrong” or the private sector. The more the Fed gets things wrong, the harder it is for the private sector to infer something about the economy from Fed behavior. This is, of course, one of the reasons behind arguments for transparency in monetary policy.

METHODOLOGY: A DÉJÀ VU EXPERIENCE

Bob Arnold’s paper does a solid job of explaining the CBO’s methodology for measuring and projecting potential output. He also shows substantial awareness of the limitations of their approach; there is little for me to add on that score. To provide a different perspective, in this section I offer readers a “blast from the past,” from 30 years ago, in fact. I describe the approach of Perloff and Wachter (1979) from a Carnegie-Rochester conference in 1978. Like Arnold, Perloff and Wachter start with an estimate of the non-accelerating inflation rate of unemployment (NAIRU) from a previous paper; then, they estimate potential labor input as follows:

$$(2) \log(n) = c + \alpha(u - u^*) + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \varepsilon_t,$$

where t^k , $k = 1, 2, 3$ are polynomial time trends, u is the unemployment rate, c is a constant, and ε is a residual. Potential labor input, n^* , is evaluated using this equation by setting cyclical and noise terms to zero; in this instance, $u = u^*$ and $\varepsilon = 0$ for all t . Perloff and Wachter follow the same procedure with potential capital input, except that the equation in this case is a “cyclically sensitive translog production function” (p. 122) augmented with more polynomial time trends. The similarity to Arnold’s equation (1) is remarkable.⁵

With this sameness in mind, I can make my job as discussant easier by shamelessly stealing from Perloff and Wachter’s discussants. Gordon (1979) focused on estimation:

[W]ithout making any statistically significant difference in the wage equation, one could come up with an estimated increase in u^* between 1956 and 1974 ranging anywhere from 0.58 to 1.61 percentage points... (p. 190)

In other words, taking u^* as exogenous, rather than estimating a complete system, particularly while ignoring the imprecision of the first-stage

⁵ From a real-time perspective, the CBO’s methodology could be more problematic than Perloff and Wachter’s in that the CBO uses trends dated back from the previous business cycle peak. No doubt this is to avoid the political heat that might come from making a call on a potentially contentious issue in real time. By definition, this method will miss turning points, possibly by wide margins.

estimates, is problematic. Elsewhere, Gordon remarks on overparameterization:

Taking this set of data for u^* , one can compute an acceptable and consistent natural output series without any use of production functions at all. (p. 188)

That is, because the time-trend variables are doing the bulk of the work, it is not clear that there is anything unambiguously “supply side” in the calculation. The other discussants, Plosser and Schwert (1979), focused on interpretation of the results and the related issue of econometric identification:

[A]ggregate demand policies are not necessarily appropriate in a world where actual output is viewed as the outcome of aggregate supply and demand...In such an equilibrium world, “potential output” ceases to have any significance. (p. 184)

Thus, even though the real business cycle literature had yet to emerge, the seeds of the idea were clearly already planted.

Both commentaries remark, in their own way, on econometric identification. How does one differentiate between supply (or potential) and demand (the gap)? Does it even make sense to try? The use of time trends, which are both *deterministic* and *smooth*, is an identifying assumption made by both Perloff and Wachter (1979) and Arnold (2009). Their use implies that supply shocks have not happened often historically and can be safely ignored in real time for forecast purposes. When Perloff and Wachter were writing, the literature on unit roots in real GDP—which would come to include, as it happens, an important contribution by Nelson and Plosser (1982)—had not yet arisen. But this is not so for the CBO or any of a variety of other institutions that use similar approaches.⁶ Why, then, has the methodology on measuring potential output apparently

not absorbed anything from the literature on unit roots and stochastic trends over the past 30 years?

My conjecture is threefold. First, the CBO—like most macroeconomic policy institutions—maintains a distinctly Keynesian perspective on how the economy works, a view that maintains that the majority of fluctuations in real GDP come from demand disturbances and that policy plays a key role in smoothing those fluctuations. This approach is natural enough; policy institutions do tend to draw individuals who believe that policy is highly consequential. And to paraphrase the old line: When one likes to use hammers, the object of interest tends to look like a nail. My second conjecture is more subtle. Economists at institutions like the CBO must be able to answer a wide variety of questions from decisionmakers and they need a structure that allows them to do so in short order. The complex, deterministic accounting structure that Arnold describes allows the CBO to do that, although one could of course quarrel with the efficacy of the advice that comes from such a structure. Third, while I would argue that the literature on unit roots shows that permanent shocks to GDP—shocks that can fairly be characterized as supply shocks—are important, that literature has not yet provided high-precision tools for measuring those shocks in real time. The standard errors of estimates of potential output and the output gap are large.⁷ And the problem gets worse as the parameter space of the model grows.

Nonetheless, I would argue that even though adopting the stochastic approach involves tackling some difficult issues, it is still a step worth taking. These same issues exist with the extant method, but they have been swept under the rug through the identification by assumption implicit in the use of time trends to represent aggregate supply. We are dealing with unobserved variables here; it only makes sense that, with the passage of time, our backcasts of potential output would differ significantly from our nowcasts. To “assume away” the stochastic properties of the data only

⁶ Barnett, Kozicki, and Petrinec (2009) note that the Bank of Canada has used a stochastic method for measuring potential since 1992. The Federal Reserve Board’s FRB/U.S. model forecast uses a stochastic state-space method. The Fed’s official *Greenbook* forecast—being judgmental—is more complicated. The Board staff consult a variety of models for guidance on adjusting potential output and its constituent parts, but they do so on an ad hoc basis. There is, however, a significant smoothness prior on trend labor productivity, and hence on potential output, and a prior that Okun’s law holds fairly strongly.

⁷ The discussion slides show an example of the bootstrapped standard errors from a simple unobserved components model of potential output. These are available at the online version of this *Review* article.

ignores the issue; it doesn't solve it. A more clear-eyed view, in my opinion, is to accept the stochastic nature of potential and adjust procedures and interpretations to this reality by being prepared to adapt estimates rapidly and efficiently in real time (see, e.g., Laxton and Tetlow, 1992).

OKUN'S LAW

I have already noted the strong Keynesian prior implicit in the methods for measuring potential output at the CBO and other policy institutions. As noted, this prior is evident in the use of deterministic time trends. It is also a function of the fact that potential output—and hence output gaps—are constructed beginning with estimates of the NAIRU, and hence the unemployment gap, using Okun's law. This is illustrated in Arnold's Figure 1, which shows the CBO output gap and the unemployment gap on the same chart. The chart provides an "ocular regression" of Okun's law: The two lines are nearly on top of one another, meaning that a linear, static relationship between the two concepts fits the (constructed) data very well. In essence, this means that the output gap and the unemployment gap are nearly the same.

The view that the unemployment gap and output gap are isomorphic—that is, the view that Okun's law really is something that approaches a "law"—has important implications for the characterization of business cycles. The following log-linearized Cobb-Douglas production function shows this:

$$(3) \quad y = a + \theta n + (1 - \theta)k,$$

where a is total factor productivity, and we measure potential output using full-employment labor input, n^* , and the actual capital stock, k , as is usually the case:

$$(4) \quad y^* = a^* + \theta n^* + (1 - \theta)k,$$

and then subtract equation (4) from equation (3) to show the relationship between output gaps, $y - y^*$, and the labor market gap, $n - n^*$ ⁸:

$$(5) \quad y - y^* = (a - a^*) + \theta(n - n^*).$$

Now Arnold's Figure 1 implies that $y - y^* - \theta(n - n^*)$ is small and unimportant—taken to the limit, Okun's law implies that it should be white noise. This, in turn, means that what we might call the *productivity gap*, $a - a^*$, must also be small and unimportant. Should it be? Should anyone care? What is the productivity gap anyway? The productivity gap can represent any or all of a variable workweek of capital, variable capacity utilization, or labor adjustment costs to productivity shocks.⁹ Loosely speaking, fluctuations in a that are not in response to shocks to a^* are labor adjustment shocks, whereas shocks to a^* , all else equal, are classic productivity shocks. The productivity gap, $(a - a^*)$, can be unimportant only in the unlikely circumstance that actual productivity, a , moves instantaneously with a productivity shock, a^* , and disturbances to a , holding a^* constant, are themselves close to white noise. In short, the only way the productivity gap could be small and unimportant—and, therefore, the only way that Okun's law can hold so tightly as to be called a law—is either because aggregate demand moves instantaneously with productivity or if there are no productivity shocks in the first place. Neither of these possibilities seems plausible.

My own preference would be to drop the deterministic time trends, relaxing somewhat the iron grip of Okun's law, and treat potential output as a stochastic variable. Doing so would allow for meaningful supply-side shocks, modeled using state-space techniques, probably with the Kalman filter.¹⁰ From an operational point of view, this shifts the prior on the incidence of shocks somewhat. Under the deterministic prior, all real surprises are demand shocks and this view

⁸ I am blurring the distinction between the unemployment gap and the labor market gap—the difference being what might be called the average workweek gap and the labor force participation rate gap. This distinction is important to my point only if one thinks that *all* productivity adjustment— a movements relative to a^* —is carried out on these two margins, which seems unlikely.

⁹ Whether there is any meaningful distinction among these three stories depends on the underlying model.

¹⁰ My suggestions here are particularly relevant for a decisionmaking body when the level of the gap is important. I think this is true for almost all policy institutions but is undoubtedly "more true" for, say, a central bank, than for a fiscal authority.

is adjusted only rarely and after the fact; with the stochastic view, the default option becomes one wherein some portion of a given output surprise is characterized as a supply shock. The model user could override that prior, but it would be a conscious decision on the user's part to do so. In this way, the stochastic approach would be responsive in real time, allowing estimates to adapt to developments such as the productivity boom of the late 1990s in a way that the deterministic approach would not. Such a property is an important one, particularly for institutions whose policy instruments may be adjusted with relatively high frequency. State-space models also allow the modeling of nonlinearities—for example, to capture different dynamics when cycles are being driven largely by supply shocks rather than by demand shocks or to allow for “jobless recoveries”—although the econometric hurdles are correspondingly higher.¹¹

Such an approach comes at some cost, however, because either the parameter space must be small or the user must be willing to impose priors on enough parameters to give the estimator a chance of producing reasonable results. Still, this approach would likely impose fewer restrictions than the current approach. At a minimum, weakening the prior that all shocks are demand shocks opens the door for model users to consider what kind of shocks might have produced the cross section of measured surprises—positive for output and negative for inflation, for example—in real time. This, in turn, would allow a more rapid adjustment to new information and smaller and less persistent forecast and policy errors than would otherwise be the case.

CONCLUSION

Bob Arnold has outlined a detailed and sophisticated approach to measuring potential output as used by the CBO. In my opinion, the approach is representative of the perspective and

needs of a range of policy institutions. In general terms, the remarkable thing about the CBO method, and methods like it, is how little it differs from methods used 30 years ago. This lack of penetration of academic ideas into the policymaking sphere is perplexing in some ways. However, it reflects, in part, the needs of institutions to be able to answer myriad questions using the same model. This practice tends to result in the construction of large, elaborate models, and unfortunately not all modern econometric techniques scale up well to large models. The good news is that new methods in Bayesian econometrics offer considerable help in estimating larger systems while paying proper heed to the priors of the model builders and users. Another source of the lack of progress, in my view, is the strong Keynesian prior regarding the sources of business cycle fluctuations. Many public policy institutions regard supply shocks as rare enough to be ignored. I would argue that this prior is overly strong—we know for a fact it was dead wrong in the United States in the late 1990s (see, e.g., Anderson and Kliesen, 2005; and Tetlow and Ironside, 2007). It might also be deleterious for policymaking because the perspective that all shocks are demand shocks leads directly to the view that all fluctuations should be smoothed out, which is arguably a recipe for “fine-tuning.”

We are now in a period in which the CBO methodology is being tested. By construction, the CBO will have concluded that the current “financial stress shock” to the U.S. economy is entirely a demand-side phenomenon with large implications for the output gap and eventually for inflation. This is a contestable position. It would not be hard to fashion an argument that the desired capital stock, and hence the level of potential output, has shifted down; interpreting the shock in this less devoutly Keynesian way would mean smaller output gaps, less disinflationary pressure, and somewhat less need for expansionary policy, all else equal. We shall see. In any case, quite apart from the methods detailed therein, Bob Arnold's paper shows a mindful understanding of the uncertainties involved, which is probably more important. It thereby serves the Congress well.

¹¹ Bayesian methods can be helpful in this regard, particularly for policy institutions that tend to be unapologetic about having prior beliefs.

REFERENCES

- Aaronson, Stephanie; Fallick, Bruce; Figura, Andrew; Pingle, Jonathan and Wascher, William. "The Recent Decline in the Labor Force Participation Rate and Its Implications for Potential Labor Supply." *Brookings Papers on Economic Activity*, Spring 2006, 1, pp. 69-134.
- Anderson, Richard G. and Kliesen, Kevin L. "Productivity Measurement and Monetary Policymaking During the 1990s." Working Paper No. 2005-067A, Federal Reserve Bank of St. Louis, October 2005; <http://research.stlouisfed.org/wp/2005/2005-067.pdf>.
- Arnold, Robert. "[The Challenges of Estimating Potential Output in Real Time](#)." Federal Reserve Bank of St. Louis *Review*, July/August 2009, 91(4), pp. 271-90.
- Barnett, Russell; Kozicki, Sharon and Petrinec, Christopher. "[Parsing Shocks: Real-Time Revisions to Gap and Growth Projections for Canada](#)." Federal Reserve Bank of St. Louis *Review*, July/August 2009, 91(4), pp. 247-65.
- Gordon, Robert. "A Comment on the Perloff and Wachter Paper." *Carnegie-Rochester Conference Series on Public Policy*, January 1979, 10(1), pp. 187-94.
- Laxton, Douglas and Tetlow, Robert J. "A Simple Multivariate Filter for the Measurement of Potential Output." Technical Report No. 59, Bank of Canada, June 1992.
- Nelson, Charles R. and Plosser, Charles I. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications." *Journal of Monetary Economics*, September 1982, 10(2), pp. 139-62.
- Perloff, Jeffrey M. and Wachter, Michael L. "A Production Function—Nonaccelerating Inflation Approach to Potential Output: Is Measured Potential Output Too High?" *Carnegie-Rochester Conference Series on Public Policy*, January 1979, 10(1), pp. 113-63.
- Plosser, Charles I. and Schwert, William G. "Potential GNP: Its Measurement and Significance: A Dissenting Opinion." *Carnegie-Rochester Conference Series on Public Policy*, January 1979, 10(1), pp. 179-86.
- Tetlow, Robert J. and Ironside, Brian. "Real-Time Model Uncertainty in the United States: The Fed, 1996-2003." *Journal of Money, Credit and Banking*, October 2007, 39(7) pp. 1533-61.